Explainable AI for Health and Well-being Risk Prediction

This project develops interpretable machine learning systems for predicting a wide range of health and well-being outcomes using structured and tabular datasets. Students will build end-to-end ML pipelines using XGBoost, Explainable Boosting Machines (EBMs), Neural Additive Models (NAMs), and modern tabular foundation models. The project examines conventional health indicators as well as psychological, social, economic, behavioral, and environmental dimensions of well-being, integrating a wide variety of measures (e.g., life satisfaction, resilience, social support, financial strain, sleep quality, physical activity, commute burden, crime, neighborhood safety). Students will design robust preprocessing workflows, train and evaluate tabular models, and apply interpretability methods such as SHAP, LIME, and surrogate modeling to uncover how demographic, social, behavioral, and environmental factors shape well-being risks. The goal is to produce transparent, interpretable models that reveal meaningful patterns and support evidence-based policy and personalized interventions.

The ideal candidate has strong Python skills and an interest in applied machine learning. Experience with supervised learning or explainable AI is helpful but not required.

Supervisor: Prof. Nicholas Spence

Application: Please submit your CV, unofficial transcript, and a single paragraph describing your interest in the project in an email to Prof. Nicholas Spence (nicholas.spence@utoronto.ca) and Prof. Ethan Fosse (ethan.fosse@utoronto.ca).

Start Date: Fall 2026 / Winter 2027

Research Area: Applied Machine Learning, Operations Research, Human Factors